TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 4479 - 4485

A Survey on Optimization Techniques in Voice Disorder Classification

N.A. Sheela Selvakumari[1], Dr. V.Radha[2]

[1]Assistant Professor, Department of Computer Science, Sri Krishna Arts and Science College, Coimbatore

[2]Professor &HoD, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore

Article Info
Volume 82
Page Number: 4479 - 4485
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 22 January 2020*

Abstract

Voice diseases are increasing dramatically, by unhealthy social habits and voice abuse. In this work, we investigated the optimization techniques for voice pathology and voice disorder classification system. The literature review presents a survey of the acoustic analysis for human voice disorder classification using optimization and machine learning techniques with importance to pathology detection and its classification. The voices input signal is given to numerous classifications methods, in such the simplest way the classification produces the result against the pathology voices and normal voices with relevence to the male and female human voice classification. The foremost goal of this analysis work is to provide a entire study of the most popular machine learning techniques namely, Noise Removal and Silence Removal and different filters in preprocessing, The feature extraction techniques namely, Acoustic features (signal energy, pitch, formant, jitter, shimmer), Reflection coefficients, Autocorrelation, Linear Predictive Coding (LPC), Mel-frequency cepstrum co-efficient (MFCC), Zero crossing with peak amplitude (ZCPA), Dynamic time wrapping (DTW) and Relative spectral processing (RASTA). And to classify the input voice signal with the help of classification algorithms like, Support Vector Machine (SVM) and Back Propagation Neural Network (BPNN).

Key Words :*Voice pathology, survey of existing literature, Optimization Techniques,classification methods.*

I. INTRODUCTION

In recent years, huge consideration has been given to the area of voice pathology ID and observing. In the voice pathology treatment, patients need to every now and again visit the specialist for voice treatment. However, the patients are trusting that quite a while will counsel, and are spending a ton of cash to discover the pathology in light of the fact that the specialists need to discover the issues in the vocal folds utilizing some endoscopic methodology as it were. Absolutely it is a costly just as a tedious procedure. Consequently such things made the patients inconvenience. This circumstance makes ready for the examination in finding a computerized instrument to distinguish the voice pathology. The essential motivation behind this computerized

January - February 2020
ISSN: 0193 - 4120 Page No. 4479 - 4485

instrument is to help the patients for recognizing the obsessive issues for further progress. Essentially, the voice pathology may cause because of the flaws in the discourse organs, Autism, psychological instability, hearing impedance, Paralysis or various handicaps. Clinically there is a wide scope of rules and strategies to discover the voice pathology and furthermore these techniques are abstract and obtrusive. Yet, it isn't the situation for the programmed instrument for the voice pathology.

The main aim of this review paper is to provide various methods and techniques available for pathological voice detection system. The section 2 describes the commonly used methods and techniques in voice pathology classification system. Section 3 focuses on the survey of recent methods in voice pathology classification system and section 4 concludes the review paper.

II. COMMONLY USED METHODS FOR VOICE PATHOLOGY IDENTIFICATION

Harar et al (2017) discussed the voice disorders and used Deep Neural Networks (DNN) for voice pathology detection. The research uses the German corpus Saarbruecken Voice Database with recurrent Long-Short-Term-Memory (LSTM) layers for voice pathology detection. It is found there is high accuracy with sensitivity and specificity on testing files and training data set. This system achieved an efficient time to find healthy persons and pathological patients. Finally, the development of Voice Pathology Classification system is a hard problem in the whole dataset with accuracy. This research motivates us to use Saarbruecken Voice Database for pathology detection with the various QoS parameters.

Zhang et al (2008) investigated the acoustic characteristics of pathologies. It concentrates on the vowels in terms of sustain as well as running. The Perturbation methods, SNR and nonlinear dynamic methods are used to investigate characteristics of pathologies. And classify as low-dimensional rate as the laryngeal pathologies and high-dimensional rate as normal voice. Finally, the research is performed by nonlinear dynamic analysis for sustain and running vowels in processing voice analysis. This helped to manage pathology voice detection with respect to the perturbation analysis.

Parsa et al (2001) investigated the Pathological Voice examination. It separates the ordinary and neurotic voices regarding the acoustic measures. The supported vowels and ceaseless discourse for acoustic estimations are taken. The discourse signals are separated from supported vowels and nonstop discourse, regarding the crucial recurrence, sufficiency irritation, ghastly measures, and glottal clamor. The outcome appears, ordinary voice, and neurotic voice characterizations are finished utilizing the acoustic measures with the higher disconnection and order on nonstop discourse. This examination inspired us for the Pathological Voice examination on the acoustic measures.

Saeedi et al (2011) proposed the Digital examination of neurotic voices. The variations from the norm of the vocal system are researched. The non-obtrusive device is utilized to deduct the uniformity of the voice by utilizing the wavelet-based technique. These techniques are creating the conclusive outcome of typical and confused voices. The Support Vector Machine calculation is utilized for highlight extraction and order on Orthogonal channel banks. The last arrangement rate is creating the typical and obsessive voices on the databases. Additionally, the hereditary calculation is executed to look for the best grouping rate. At long last, the proposed calculation can arrange the scattered sign with a excellent choice for pathological voices.

Fezari et al (2014) talked about the acoustic voice examination for undesirable voice infections voice misuse conclusion. And furthermore examine the voice pathologies location voice symptomatic. The solid and obsessive voices, non neurological from the German database which contains numerous sicknesses are grouped. Likewise, the directed calculation is utilized for voice pathologies

January - February 2020
ISSN: 0193 - 4120 Page No. 4479 - 4485

identification. To achieve this undertaking, Mel Frequency Cepstral Coefficients are displayed by a weighted Gaussian Mixture Model (GMM) and plotted for highlights extraction. This obsessive voices discovery utilizes the acoustic voice examination, with programmed speaker acknowledgment procedures. This finds the MFCC, frequencies, and amplitudes. This groups the exact obsessive voices. At last, this examination identifies typical and strange voices regarding the MFCC and Energy-Optimal framework.

Villa-Canas et al (2012) examined the practical issue and laryngeal pathologies. Broke down those clutters by utilizing vibrational examples of the vocal folds. The non-parametric cepstral coefficients in Mel channel bank scales were utilized for programmed include choice. The significant highlights are naturally chosen utilizing Principal Components Analysis (PCA) and Sequential Floating Features Selection (SFFS). Voice is solid or scattered was chosen, with the assistance of direct and quadratic Bayesian, K closest neighbors and Parzen strategies.

Zhang et al (2017) proposed the automatic and rapid computer-aided diagnosis system for pathological brain images detection. Magnetic Resonance Imaging (MRI) is used for simplification and classification. The classification result might be pathological or normal. For feature extraction, the Hue Moment Invariants (HMI), and for classification the used of Twin Support Vector Machine (TSVM) and Generalized Eigen value Proximal SVM (GEPSVM) are used. Finally, the methods “HMI + GEPSVM” and “HMI + TSVM” achievedhigherclassification accuracy on pathological brain detection and are proposed. These contributionsare used to concentrate the better classification performance.

Zhang et al (2017) proposed the MDNet to coordinate multimodal mapping. This maps therapeutic pictures and diagnostic reports. The indicative reports are finished by envisioning the consideration. The restorative pictures mapping is finished by utilizing the manifestation depictions. That technique analysis is utilized to understand pictures, produce indicative reports and recover pictures. This examination is utilized to upgrade the element extraction and productivity in an advanced way. This examination makes the multimodal mapping idea to oversee voice in order in various criteria.

Wang et al (2017) proposed the Optimization systems for Pathological Brain Detection. The by means of Wavelet Packet Tsallis Entropy for Detection and Real-Coded Biogeography for Optimization are utilized. This exploration is recognizing obsessive minds with the assistance of Novel Pathological Brain Detection (PBD) regarding the Feed Forward Neural Network (FNN). This examination shows the proposed WPTE + FNN + RCBBO delivered precision superior to the existingapproach. This exploration proposes the Feed Forward Neural Network (FNN) approach for conclusive streamlining in the voice pathology location.

Wang et al (2017) assessed the back propagation neural system calculation on the restorative instrument of Amomumcompactumgentamicin. Intense Kidney Injury (AKI) on the Back Propagation Neural Network (BPNN) is taken. These models were set up for characterizing information from the control, model, and AC-treated gatherings. At long last, the BP neural system calculation was used for the grouping of information, and the exactness rates for order were great in both positive and negative spectra. This examination has helped us to acquire a superior comprehension of protective mechanisms of order just as the Backpropagation Neural Network calculation.

Kukharchik et al (2007) proposed the vocal pathology diagnostic method. The acoustic examination comprises of wavelets and pseudo-wavelets for include extraction. In addition,

January - February 2020
ISSN: 0193 - 4120 Page No. 4479 - 4485

ceaseless wavelet and wavelet-like change are utilized for vocal pathology indicative techniques. The immediate vocal overlay perception characterizes the discourse signal with the assistance of SVMs and Neural Networks. In the order technique, Support vector machine is utilized in crease pathology identification. Quiet Removal and Vowel Extraction were additionally used to improve order results with higher precision.

Saeedi et al (2011) analyzed the obsessive voice investigation process. Non-obtrusive instrument for the location is utilized as the fundamental objective of neurotic voice examination. The typical and confused voices are grouped dependent on the wavelet-based technique. The Support Vector Machines are joined with Wavelet channel banks. This blend is utilized infeature extractors and classifiers. The Support Vector Machines are utilized for arrangement and Wavelet channel for highlight extraction. The Orthogonal channel (cross section) with a hereditary calculation with a wellness work is proposed to adjust the grouping procedure in a productive way. This proposed technique corrects the typical and neurotic voice signals with a higher order rate on databases.

Saidi et al (2015) proposed the non-intrusive technique. This procedure is utilized to arrange normal voice flag and confusion voice signals. The wavelet disintegration is utilized for include extraction with better recurrence. Moreover, the M-band wavelet decay is additionally applied for highlight extraction. In the wake of applying this calculation, a hereditary calculation is utilized for ideal wavelet separating. Bolster Vector Machine is utilized in the last grouping of typical and obsessive voices in the database. To deal with the characterization procedure, the noninvasive strategy in particular M-band wavelet framework, five-band wavelets, and hereditary calculation are suggested for improved arrangement execution.

Peng et al (2007) discussed about the characterization of pathological voice. The 30acoustic highlights are considered to group solid and irregular voices. The PCA involved the element space change and information dimension reduction. The Support Vector Machine (SVM) calculation is utilized for arrangement of healthy and neurotic voices. At long last, the discovery and arrangement rates are estimated as for Sensitivity and Specificity and Feasibility of voice signal order. The component choice is finished by utilizing LSSVM classifier with the RBF portion work.

III. SURVEY OF METHODS USED IN VOICE PATHOLOGY IDENTIFICATION

S.No	Author & Year	Title of the Paper	Methods
1	Ali, Z., Elamvazuthi, I. and Alsulaiman, M.	Detection of Voice Pathology using Fractal Dimension in a Multi resolution Analysis of Normal and Disordered Speech Signals	Voice pathology detection – Wavelet transformation & Fractal dimension - Katz algorithm – Higuchi algorithm
2	ImanEsmailia,NaderJafarniaDabanlooa, Mansour Vali	Automatic classification of speech dysfluencies in continuous speech based on similarity measures and morphologi	Automatic dysfluency classification - cross-correlation and Euclidian similari

January - February 2020
ISSN: 0193 - 4120 Page No. 4479 - 4485

		cal image processing tools	ty measures-Morphological processing tools
3	ZhijianWang,Ping Yu, Nan Yan, Lan Wang &Manwa L. Ng	Automatic Assessment of Pathological Voice Quality Using Multidimensional Acoustic Analysis Based on the GRBAS Scale	Automatic Voice quality assessment - Acoustic analysis - GRBAS - SVM & Extreme machine learning
4	DariaHemmerlingn, AndrzejSkalski,JanuszGajda	Voice data mining for laryngeal pathology assessment	Voice pathology detection- Acoustic analysis - Random forest
5	PouriaSaidi, FarshadAlmasganj	Voice Disorder Signal Classification Using M-Band Wavelets and Support Vector Machine	Voice disorder classification - M-band wavelets - Support vector machine & Genetic algorithm
6	Virgilijus Uloza, EvaldasPadervinkis	Exploring the feasibility of smart phone microphone for measurement of acoustic voice parameters and voice pathology screening	Smart Phone voice recordings-frequency,jitter,shimmer and NNE-classification CCR, RFC, EER
7	Ghulam Muhammad, MehediMasud	Spectro-temporal directional derivative based automatic speech recognition for a serious game scenario	ASR system-STDD – Discrete Cosine Transform
8	Jus hen, Chang peng Ti	Automatic video self modeling for voice disorder	Video self modeling-video similarity metrics-lip synchronization-re-sampling scheme
9	Bandini. B,	Automatic identificatio	Parkinson's

January - February 2020
ISSN: 0193 - 4120 Page No. 4479 - 4485

	Giovannelli. F	n of dysprosody in idiopathic parkinson's disease	Disease - Dysprosody- Objective Parameters – Automatic Acoustic Analysis
10	Thuy Ngoc Tran, William Cowley	Automatic adaptive speech separation using beam former-output-ratio for voice activity classification	Beamforming-BiBeam-Speech separation-Voice activity classification-Beamformer output ratio-Adaptation control

From the above table, one can induce that the greater part of the current voice pathology order frameworks utilizes the acoustic highlights. At the point when we utilizing different acoustic highlight which best portray the working and state of different discourse organs to break down the voice change during pubescence period. Pitch is a trait which speaks to the structure and size of the larynx and vocal folds. The voice quality can be estimated utilizing GRBAS scale, with the goal that the voice quality will be reviewed likewise. The AI calculations are utilized broadly so as to group the ordinary and neurotic voices. Mel Frequency Cepstral Coefficient is most normally utilized in discourse acknowledgment framework. Zero intersection top sufficiency is regularly utilized in programmed discourse acknowledgment framework in uproarious situations (Kim, J., et al 2015). DTW is usually utilized for estimating likeness between two fleeting arrangements which may fluctuate in time or speed (Arias-Londoño, J. D., et al 2011). RASTA strategy is commonly utilized for discourse investigation in which discourse signals are improved to build up a commotion hearty discourse acknowledgment framework.

CONCLUSION

The methodologies accessible in voice pathology discovery process for highlight extraction, include choice, and voice pathology order strategies are investigated. Be that as it may, the typical and pathology arrangement is a difficult undertaking against the two unique informational index to be specific Saarbruecken informational collection and possess Real-Time dataset, especially in Saarbruecken informational index, the voice pathology discovery can't be precisely grouped with the ordinary voice signal. Related works are talked about the novel strategies which arrange all information concerning the conveyance over a pathology grouping. This audit paper is centered around the strategies for voice pathology and voice ailment grouping framework. It gives the detail learn about the strategies and various methods for voice issue order. The primary point of this examination work is to comprehend the procedure of the strategies to build up an effective voice pathology framework with extraordinary precision. This audit paper will help the specialists ready to work in the territory of voice pathology characterization framework with acoustic highlights.

REFERENCES

[1] Arias-Londoño, J. D., Godino-Llorente, J. I., Markaki, M., & Stylianou, Y. (2011). "On combining information from modulation spectra and melfrequency cepstral coefficients for

January - February 2020
ISSN: 0193 - 4120 Page No. 4479 - 4485

automatic detection of pathological voices". Logopedics Phoniatrics Vocology, 36(2), 60-69.

[2] Fezari, M., Amara, F., & El-Emary, I. M. (2014), "Acoustic analysis for detection of voice disorders using adaptive features and classifiers", In Proc. 2014th Int. Conf. on Circuits, Systems, and Control.

[3] Harar, P., Alonso-Hernandezy, J. B., Mekyska, J., Galaz, Z.,Burget, R., &Smekal, Z.(2017), "Voice pathology detection using deep learning: a preliminary study", In 2017 international conference and workshop on bioinspired intelligence (IWOBI), Pp: 1-4.

[4] Kim, J., Kumar, N., Tsiartas, A., Li, M., & Narayanan, S. S. (2015). Automatic intelligibility classification of sentence-level pathological speech.Computer speech & language, 29(1), 132144.

[5] Kukharchik, P., Martynov, D, Kheidorov, I., &Kotov, O. (2007), "Vocal fold pathology detection using modified wavelet-like features and support vector machines", In 2007 15th European Signal Processing Conference,Pp: 2214-2218.

[6] Parsa, V., & Jamieson, D. G. (2001),"Acoustic discrimination of pathological voice: sustained vowels versus continuous speech", Journal of Speech, Language, and Hearing Research, Vol. 44(2), Pp: 327-339.

[7] Peng, C., Chen, W., Zhu, X., Wan, B., & Wei, D. (2007), "Pathological voice classification based on a single Vowel's acoustic features", In 7th IEEE International Conference on Computer and Information Technology (CIT 2007), Pp: 1106-1110.

[8] Saeedi, N. E., Almasganj, F., &Torabinejad, F. (2011),"Support vector wavelet adaptation for pathological voice assessment", Computers in biology and medicine, Vol. 41(9), Pp: 822-828.

[9] Saidi, P., &Almasganj, F. (2015), "Voice disorder signal classification using m-band wavelets and support vector machine", Circuits, Systems, and Signal Processing, Vol. 34(8), Pp: 2727-2738.

[10] Villa-Canas, T., Belalcazar-Bolaños, E., Bedoya-Jaramillo, S., Garces, J. F., Orozco-Arroyave, J. R., Arias-Londono, J. D., & Vargas-Bonilla, J. F. (2012), "Automatic detection of laryngeal pathologies using cepstral analysis in Mel and Bark scales", In 2012 XVII Symposium of Image, Signal Processing, and Artificial Vision (STSIVA) , Pp: 116-121.

[11] Wang, S., Du, S., Li, Y., Lu, H., Yang, M., Liu, B., & Zhang, Y. (2017), "Hearing loss detection in medical multimedia data by discrete wavelet packet entropy and single-hidden layer neuralnetwork trained by adaptive learning-rate backpropagation.", In International Symposium on Neural Networks, Springer, Cham, Pp: 541-549.

[12] Wang, S., Li, P., Chen, P., Phillips, P., Liu, G., Du, S., & Zhang, Y. (2017), "Pathological brain detection via wavelet packet tsallis entropy and real-coded biogeography-based optimization.", Fundamenta Informatica, Vol. 151(1-4), Pp: 275-291.

[13] Zhang, Y., & Jiang, J. J. (2008), "Acoustic analyses of sustained and running voices from patients with laryngeal pathologies", Journal of Voice, Vol. 22(1), Pp: 1-9.

[14] Zhang, Y., Yang, J., Wang, S., Dong, Z., & Phillips, P. (2017), "Pathological brain detection in MRI scanning via Hu moment invariants and machine learning", Journal of Experimental & Theoretical Artificial Intelligence, Vol. 29(2), Pp: 299-312.

[15] Zhang, Z., Xie, Y., Xing, F.,McGough, M., & Yang, L. (2017), "Mazenet: A semantically and visually interpretable medical image diagnosis network", In Proceedings of the IEEE conference on computer vision and pattern recognition, Pp: 6428-6436.